

効果的なテストの作成

—生徒の力を伸ばすテストで、より良い授業を—

佐藤 臨太郎

1. はじめに

リーディング、リスニング、ライティング、スピーキングの4技能をバランスよく測るために、従来のセンター試験に代わり、外部テストの導入がほぼ確実になるなど、入試英語は大改革の方向に進んでいます。これは入試を変えることにより、授業改革を促し、日本の英語教育をより良い方向へと導くとの意図からです。「試験が変われば授業が変わる」「テストのために授業を行う」。このことに違和感を持つ先生もおられるかもしれません。確かに、試験を変えることにより英語教育もすぐ変わるというような単純なことではありません。しかしながら、生活の中で英語を使う場面がほとんどない EFL (English as a foreign language) 環境において、多くの学習者が学校科目のひとつとして、入試で重要な「英語」をテストのために勉強しているということも事実ではないでしょうか。そうであるならば、学習者の4技能における英語力をより正確に測ることのできる効果的なテストを作成することにより、授業の質を向上させ、生徒の英語力を伸ばしていくことが重要であるといえます。本稿では、どのように、生徒の力を伸ばすことの出来る効果的なテストを作成すべきかについて議論していきます。

2. テストの基礎概念について

テスト作成に当たって重要な観点として妥当性、信頼性、実用性、波及効果があります。

妥当性とは、そのテストが測定しようとしている能力をしっかりと測っているかということです。例えば、校内陸上競技大会でのクラスのリレー代表選手を決めるのに、たまたまその日に水泳の授業があるので、自由形のタイムを計って決定するということはありませんよね。短距離走と水泳の自由形では使う能力が違うからです。これは極端な例ですが、英語のテストでは、ペーパーテストで単語の強いア

クセントの位置を問う問題はどうか。この問題が実際にその単語を発音できるかどうかを測る目的で出題されたとする、妥当性は低いと言わざるを得ません。単語についての知識を知っていることと、発音できることは別の能力であると考えられるからです(若林・根岸, 1993)。また、このようなテストを出題し続けると、生徒は発音の練習はせず、発音についての知識を覚えるための勉強をするということも考えられます。このテストが学習者の勉強方法や授業に与える影響を波及効果といい、マイナスとプラスの両面があります。例にあげた発音問題はマイナスの波及効果となり、他には、試験で英文和訳を課すと、英語は日本語に訳して読むという態度を育ててしまう可能性がある(静, 2002)などがあげられます。プラスの波及効果としては、例えば、センター試験でリスニングテストが導入されてから、実際に高校の授業でのリスニングも重要視されるようになったことがあげられます。

さて、妥当性と同時に重要な観点が信頼性です。これはそのテストが、学習者の能力をしっかりと引き出しているか、測定の精度がどの程度高いかということです。問題の不備により、生徒に能力はあるのに発揮できないということは避けなければいけません。具体的な工夫については後で述べます。この妥当性と信頼性の関係ですが、まずは妥当性が高くなければテストとしては問題があるということです。先の発音問題ですが、発音についての知識のある生徒はコンスタントに高い点数を取り続けると考えられる(能力を発揮する)ので、テストとしての信頼性は高いかもしれません。しかしながら、もともと測ろうとする能力を測れていないので、テスト結果は誤った情報を提供しているということになりかねません。授業ではタスクやコミュニケーション活動を多く行いながら、定期考査では読解や文法問題を問う場合も、信頼性は高そうですが妥当性は低いとい

えます。実際の英語のテストでは、妥当性を若干、犠牲にしながらも信頼性の高い問題を採用しているという現状がある(Alderson 他, 1995)のですが、第一に妥当性を考慮し、信頼性の高い問題作成していくことが必要です。妥当性と信頼性の双方を完璧に求めることは現実的には不可能ですが、どこかで折り合いをつけ、できるだけ両方をバランスよく上げていく必要があるといえます。

信頼性、妥当性がともに高く、プラスの波及効果のあるテストを実施するには、膨大な時間や労力がかかることがあります。例えば、スピーキング能力を測るための口頭面談テストを3クラス、合計120名で行う場合を考えてみると、ひとり10分として、合計20時間必要です。さらに、録音したやり取りをすべて聞いて点数をつけていくためには膨大な時間がかかり、現実的には非常に難しいと言えます。このように、テストが実際に実行可能かどうかという概念や程度を実用性と言います。こども折り合いをつけて考えることが必要です。

3. 妥当性、信頼性を考慮した考査づくり

妥当性と信頼性を高めるうえで重要なポイントについて笠原・佐藤(2017)を基にして、一部を抜粋しながら紹介していきます。

① 指示を明確にする

以下は、実際にある中学校で出題された問題です。

空欄に英単語を入れなさい。

Otani is a () player.

正解は baseball とのことですが、他にも多数の正解が考えられます(例 nice, good, pro, tall …)。スポーツの名前を授業で扱った後の小テストなので、出題者は当然、生徒は baseball を入れると考えたのでしょうか、生徒の受け止め方は必ずしもそうとは限りません。また、すべての生徒が大谷選手を知っているとは限りません。残念ながら、この問題は妥当性も信頼性も低いと言わざるをえません。修正案としては、単に「以下の競技名を英語で書きなさい。(1)野球 (2)テニス…」や、授業において先生が大谷選手について話題としているという前提ならば、「大谷選手がプレーしているスポーツ名を英語で書

きなさい」などが考えられます。このように、問題を作る際は、指示を明確に、作題意図との受験者の受け取り方とのギャップのないようにしなければいけません。以下の3点がポイントです。

- ・問う知識、測る能力を明確にする
- ・評価の規準、観点を明確にする
- ・測るべきことが測れるように問題を作成し、実施、採点する。

② できるだけ直接的に測定したい能力を測る

前節でも触れましたが、ペーパーテストで発音に関する知識を問う問題は、テストとして疑問です。正解できることが発音できることを(必ずしも)意味しないからです。生徒の発音を評価したいのであれば、実際に発音させること。つまり、直接、測りたい能力を測ることが理想です。英作文能力を測りたいのであれば、単語を並べかえて英文を作らせる並べ替え問題や、文法の知識を問う問題ではなく、実際にまとまった英文を書かせる。スピーキング能力を見たいのであれば、実際に会話をさせてみるなど、直接テストするのが望ましいといえます。しかしながら、実用性の面で困難が伴うことも事実です。[出来る限り]を意識して、難しいのであれば、より妥当性・信頼性の高い間接テストを作成するよう心掛けなければいけません。

③ 何を測るのか明確にする

以下の問題を見てください。

次の英文は達也が、英語の授業でマレーシアでのボランティア活動について書いた学級新聞の一部です。これを読んで間に答えなさい。

This summer I went to Malaysia to try a tree-planting volunteer activity and I had a wonderful time there. I'll never forget the experience I () Malaysia this summer.

本文の内容に合うように()に2語の単語を入れなさい。(2014年度北海道公立高校入試問題)

正解 had in

この問題は、空欄のある英文を読むだけで解答する

ことができ、他の英文を読む必要がありません。「内容に合うように」と英文理解を問うのであれば、以下のような修正が考えられます。

本文の内容に合うように、記事の空欄に入る語を選びなさい。

This summer I went to Malaysia to try a tree-planting volunteer activity and I had a wonderful time there. I will not () the experience I had in Malaysia this summer.

a. forget b. remember c. repeat d. tell

これだと、前文の理解を前提として正解を選ぶことになります。このように、問題の作成に当たっては、どの能力を問う問題になっているのかをしっかりと確認する必要があります。

④ 項目と項目が独立するようにする

以下の問題を見てください。

次の文を〔 〕内の指示に従って書き換えなさい。

- (1) Tom sings very well. [過去進行形の文に]
 (2) [(1)の解答の文を否定文に]
 (3) We will play rugby tomorrow.
 [willを使った未来を表す文に]
 (4) [(3)の解答の文を否定文に]
 (5) Linda cleans her room in the mornings.
 [「～するべきである」を表す文に]
 (6) [(7)の解答の文を否定文に]

この例では、(2)(4)(6)を正解するためには、(1)(3)(5)が正解していることが前提となります。つまり、(1)(3)(5)を間違えた生徒は、自動的に(2)(4)(6)も間違えることになり、信頼性の上で問題があります。解決策として、あまり強く関連する知識を再度問わない。つまり、(1)(2)、(3)(4)、(5)(6)のそれぞれどちらかにする、あるいは修正案として、各問を独立させるために、以下が考えられます。

- (2) Ken was playing the guitar in his room.
 [否定文に]
 (4) They will go to Hokkaido this summer.
 [否定文に]
 (6) We should eat a lot at night. [否定文に]

not を付けて否定形にすればよだけという点と、危惧としては、(2)(4)(6)がそれぞれ(1)(3)(5)を解くヒントになるということが考えられますが、定期考査においては、例えば、(2)を見て過去進行形に気づく(思い出す)のは、考査自体が学習機会であるという立場をとり許容したいと考えます。各項目を独立させ、生徒が前の問題が出来なかったことを引きずらないで、新たな問題で“フレッシュスタート”させてあげることが大切です。

⑤ 並べ替え/整序問題に関して

以下は単語を並べ替え英文を作り、記号を答える問題です。

A : I like visiting Kyoto.

B : () () (A) () (B) () there?

A : Three times. It's a very historical place, so there are still a lot of things to see.

[注] historical : 歴史的な

ア have イ many ウ you エ been オ how
カ times

(2016年度兵庫県公立高校入試問題)

答えは how many times have you been の順番となり、A が times、B が you となり、カとウを入れて正解となり得点が得られます。この並べ替えテストの問題点は、若林・根岸(1993)で20年以上前にすでに指摘されているように、3番目と5番目の語が正しいということが、他の語の位置もすべて正しいということを必ずしも意味しないということです。偶然、2語の位置だけが当たり、得点するということがあり得ます。さらにこの手の問題を解答するときには、おそらく受験生は、問題用紙のオ how イ many カ times ア have ウ you エ been に順番を振り、3番目と5番目に来るカ、ウを解答用紙に記入することになるのですが、その際に記号と何番目に入れるのかの整合を不注意で間違ってしまう可能性も考えられます。例えば、不注意で、Aに2番目や4番目に来る語を入れてしまったり、Bに4番目に来る語を入れてしまったりするケースです。その生徒の責任だと言ってしまうとその通りかもしれませんが、出来るだけそのような危惧を排除すべきかと考えます。結論から言うと、

このタイプの問題は信頼性、妥当性ともに高くありません。実際に定期考査で出題する際には、並べ替えて完成した英文、上の例だと、(How) (many) (times) (have) (you) (been) there? を答えさせる形式にするのが良いでしょう。

⑥ 同じ意味(内容)選択肢を避ける

選択肢の中から正解を選ばせる多肢選択問題を作成するにあたって、注意が必要です。次の選択肢を見てください。

- 問 次のうち本文の内容に合う英文を選びなさい。
- ア Keisuke did not play soccer when he was in America.
- イ Keisuke enjoyed playing soccer very much in America.
- ウ Keisuke had troubles when he played soccer in America.
- エ Keisuke was very happy that he played soccer in America.

アだけ「サッカーをしていない」ということで、他とは前提が違っているので正解ではなさそうです。そこで、残りの3つですが、イとエはほぼ同じことを述べています。1つを選択するわけですから、この2つは正解の可能性がなくなり、従って残ったウが正解となります。

問 空欄に入る語を選びなさい

He asked me () my parents lived.
ア if イ what ウ where エ whether

この問題もアとエが「～どうか」と同じ意味を表すので自動的に正解から外れ、実質イとウの2択の問題になっています。このように同じ意味を表す英文や語を入れると、選択肢として機能しなくなります。受験生は選択肢を見て明らかに違うものを除外し、似たような意味を持つ選択肢を避けるというテクニックを使うことがわかっています(Alderson 他, 1995)。このように、同じ意味あるいはよく似た意味を表す選択肢を入れることは避けるべきであると言えます。

(紙面の関係で5つの提案になりましたが、他については笠原・佐藤(2017)を参照してください)

4. 英文和訳について

筆者は大学入試問題にも、中学・高校における定期テストにも英文和訳問題を出題すべきではないと考えています。理由は、コミュニケーション能力を測定するツールとしては不適格であるからです。古いヒット曲で、*That's What Friends Are For* というタイトルの歌がありました。これを生徒が「それは友達がそのためにあるところのものである」と訳したとしたら、あなたはどうか判断しますか。古い訳読式の授業を受けた生徒であれば、このように訳してしまうかもしれません。果たしてこの生徒はこの英文の意味を「理解して」いるのでしょうか。少なくともこの訳では、どういうときにこの言い方が使われるのかはわからないでしょう。5点満点だとしたら何点を与えますか。判断に苦しむところです。コミュニケーション上から言えば、この言い方は「だって友達でしょう」「そのために友達がいるんじゃない」ぐらいの意味で、友人に何かをしてあげて“Thank you.”と言われたら、“You're welcome.”と言うかわりに使ってもいいよ、程度の理解があればいいところです。しかし和訳せよと言われると、受験者は本当に困ってしまいます。できる生徒なら、関係代名詞 what を理解していることを示すように訳さなければいけないのだろうか、それともわかりやすい日本語にすべきだろうか、と悩んでしまうかもしれません。できる受験者はこうした「直訳」か「意識」かの選択に悩み、できない受験者は適当に和訳をでっちあげるかもしれません。採点者はこうした訳からどれだけ受験者が理解できているかを判断するのに苦しむ上に、直訳か意識のどちらを高く評価するか、という問題にも直面します。

大学入試からも和訳問題の例を挙げてみましょう。

下線部を日本語に訳せ。

The most important difference between the way people use their kitchens today and the way previous generations used them is the range of options that make it possible not to cook. Now, many cooks rely on prepackaged frozen food for some, if not all, of their meals.

Other options include takeout form local restaurants or pizza delivery. (以下省略)
(2015年岩手大学前期日程人文社会科学部入試問題)

最初の一文を読んで筆者の頭に浮かんだのは、日本語にすれば「今の人は昔に比べて調理をしなくなった」ということです。さらに読み進めると、冷凍食品やらテイクアウトの食品が出てきたからだとわかります。そうした調理しなくても済む加工食品を選択する幅が広がったということか、それが the range of options なんだな、ということが頭の中で展開されました。ここまで十数秒くらいですが、最初の一文を訳して書くとなるとかなりの時間がかかります。直訳調でやってみると、「今の人たちと以前の世代の台所の使い方における最大の違いは、調理しないということをも可能にする選択の幅である」これだと意味不明ですね。そこでもう少し意味がわかるように意識してみると、「今の人たちがそれまでの世代と台所の使い方大きく異なっているのは、調理しなくて済むという選択の幅が広がっていることである」文脈からすると「広がっている」ということばを入れたほうがわかりやすいと思いますが、入れることで減点されるかもしれない、との考えも頭をよぎります。どこまで求められているかが不明なので、入れるべきか否か迷いました。このように、どうしたら自分の理解を示せる日本語になるか、かなりの時間を費やしてしまいます。これでは若林・根岸(1993)が言うとおりに、「英語に関するどうい能力をテストしようとしているのかわからない」ですね。例えば多肢選択式の問題や違う形式で内容理解を問うようにすれば、同じ時間内でもっと多角的に受験者の読解力を測定することができるはずで、具体的には、提示した英文の要約を問う問題、英文中から抜かれた単語をもとの位置に入れさせる問題、英文を並べ替えパッセージを完成させる問題、英文を補充させる問題、クローズテスト、Gap Filling、短答問題、情報転移等、様々な工夫が考えられます(笠原、佐藤、2017 参照)。授業において内容理解を確認するため、あるいは、構文を確認するために選択的に最小限にとどめて和訳を用いるのは有りうるかも知れませんが、考査においては、使用すべきではないと再度強調したいと思いません。

おわりに

本稿では効果的な考査を作成について、理論と実践面から述べてきましたが、紙面の関係で十分に議論できなかつたところもありますし、すぐには実現の難しい提案や、皆さんと意見の異なるところもあるかもしれません。しかしながら、本稿を読み、効果的なテストの作成についてより関心を持たれ、工夫し、試行錯誤しながら、より生徒にとってためになる考査を作成していただけたら幸いです。

参考文献

- Alderson, J. C., Clapham, C., & Wall, D. (1995). *Language test construction and evaluation*. Cambridge University Press.
- 笠原究, 佐藤臨太郎. (2017). 『英語テスト作成入門 効果的なテストで授業を変える』東京: 金星堂.
- 静哲人. (2002). 『英語テスト作成の達人マニュアル』東京: 大修館書店.
- 若林俊輔, 根岸雅史. (1993). 『無責任なテストが「落ちこぼれ」を作る』東京: 大修館書店.

(奈良教育大学 教授)